

基于视觉特征的水果蔬菜自动分类方法*

何江萍, 马彦, 李强

(兰州财经大学 信息工程学院, 兰州 730020)

摘要:为实现超市中水果蔬菜等产品的自动销售,提出了一种基于视觉特征的水果蔬菜自动分类方法。首先将所获得的水果蔬菜图像划分为多个重合的子块;接着提取这些子块的视觉特征,即尺度不变特征和方向梯度直方图特征;为了提高特征的表征能力,还将这些提取出的特征融合在一起描述目标;然后对融合后的特征做编码和池化操作以降低特征维数并提高特征区分能力;最后用所得特征训练支持向量机分类器并最终实现水果蔬菜的自动识别与分类。与现有方法相比,提出的方法在超市农产品数据库上取得了较高的识别率,从而为实现水果蔬菜的自动销售提供了技术支持和理论保障。

关键词:水果蔬菜图像;自动分类;视觉特征;编码;池化

中图分类号:TP391.4

文献标志码:A

文章编号:1672-6693(2016)03-0115-06

在城市中超市是人们获取生活物资的主要渠道。超市中商品的种类日渐繁多,实现自动区分商品可以降低管理人员的劳动强度和消费者购物付款的时间,极大地提高超市的运行效率。目前常用条形码来区分不同类的商品,但在农产品销售中条形码并不是最佳的方法,原因是大多数农产品都由很多个体构成(比如土豆、苹果等),很难为每个个体贴上条形码;此外农产品需要称重,这就要求超市销售人员必须学会辨识各类农产品并熟记其价格,因此极大地增加了超市销售的经济和时间成本。随着信息技术的发展,基于计算机视觉的农产品自动识别和分类是解决上述问题的有效途径,相应的国内外的学者也提出了许多蔬菜水果自动分类的方法。

部分相关工作主要关注对分类器的研究,比如 Anderson Rocha^[1]等人提出了一种水果蔬菜产品自动分类方法,该方法将多种特征和分类器融合在一起实现农产品自动分类。在该方法中所有的特征级联在一起然后分别训练不同的分类器并用这些分类器自动区分各种农产品。实验表明该方法具有较高的分类效果。

罗承成^[2]等人提出了一种基于多示例学习的农产品图像识别方法。该方法采用改进的单个块及其邻域提取算法将训练样本组织成多示例包,利用多样性密度算法对正包和反包进行多示例学习,根据多样性密度最大化模型对测试样本进行识别。该方法能够识别不同光照和干扰物的条件下,以任意方式摆放的多类别混合果蔬图像,且对于单类别果蔬图像的识别优于全局方法。实验表明基于多示例学习的图像识别方法具有一定的可行性。

与前述方法不同,部分相关工作主要集中于特征提取,比如 Hridkamol Biswas^[3]等人利用视觉系统的特点提出了利用计算机视觉技术实现蔬菜分类的方法。首先获取图像特征,比如颜色、形状、大小、纹理等,然后建立分类器实现蔬菜的自动分类。实验显示该方法取得了较高的分类精度,可以适用于不同场合的蔬菜分类。

陶华伟^[4]等人提出了一种基于颜色及纹理特征的果蔬种类识别方法。为更好地表述果蔬图像,首先提出一种颜色完全局部二值模式纹理特征提取图像纹理特征,然后利用 HSV 颜色直方图等提取图像颜色特征,最后采用匹配得分融合算法将颜色和纹理特征相融合表征果蔬图像。

Shiv Ram Dubey 和 Anand Singh Jalal^[5]提出了一种将多种最新的颜色与纹理特征方法结合在一起的框架,因为各类不同方法可以取长补短,所以基于这个框架提取出的特征具有很高的表述能力。实验结果显示这种将多种方法结合在一起的策略可以取得较好的识别效率。

还有部分工作同时从分类器和特征提取入手,提出了新的蔬菜水果自动分类方法。比如 Yudong Zhang^[6]

* 收稿日期:2015-07-25 修回日期:2015-10-23 网络出版时间:2016-04-29 18:34

资助项目:甘肃省高等学校基本科研业务费项目;甘肃省高等学校科研资助项目(No. 2015A-087)

作者简介:何江萍,副教授,博士,研究方向为图像处理、模式识别,E-mail: hejp_ping@qq.com

网络出版地址: <http://www.cnki.net/kcms/detail/50.1165.n.20160429.1834.002.html>

等人首先基于混沌人工蜂群算法(Fitness-Scaled chaotic artificial bee colony, FSCABC)和前馈神经网络(Feed-forward neural network, FNN)提出了一种混合分类器,接着分别提取颜色直方图、纹理和形状特征并形成新的特征,最后依据降维后的特征用提出的混合分类器对输入的水果图像做出判断。

针对农产品的自动分类已经提出了很多解决方案^[7],但由于超市中存储和销售的环境复杂多变,水果蔬菜的正确分类率仍然有待提高。基于对现有方法的研究和分析,本文提出了一种多特征提取和融合的水果蔬菜自动分类方法,为降低融合后特征维数过高而引起的“维数诅咒”现象,还利用编码(Coding)和池化(Pooling)操作对融合后的特征作进一步处理以提高特征的区分能力。

本文其余部分组织如下,在第一节中详细介绍提出的水果蔬菜自动分类方法,在第二节中将提出的方法与现有方法做了实验对比,最后在第三节总结了全文。

1 水果蔬菜自动分类方法

提出的水果蔬菜自动分类方法由两部分构成,即训练过程和分类过程,方法流程如图 1 所示。训练过程主要包括子块划分、特征提取以及编码和池化等 3 个步骤。图像子块是原始图像中的一部分,从特征表示的角度来讲图像中的子块比原始图像具有较低的维度,并且可以很好地反映图像的局部特征。

在本文方法中首先选择大小合适的窗口,然后将该窗口在图像中滑动,这样就能将原始图像划分为多个重叠的子块,最后提取每一个子块中的视觉特征。

特征提取一直是图像处理中的研究重点,在特征提取方法中尺度不变特征(Scale-invariant feature transform, SIFT)^[8]和方向梯度直方图(Histogram of oriented gradient, HOG)^[9]特征由于其独有的特性在图像处理中得到了广泛的应用。

SIFT 是一种检测局部特征的算法,该算法从一幅图中的特征点及其有关比例和方向获得了对图像的良好表征效果。与 SIFT 不同, HOG 通过计算和统计图像局部区域的梯度方向直方图来构成特征。

每一种特征提取方法都有其优势和不足,提高特征区分能力的常用的策略就是取长补短实现特征融合,因此在本文方法中首先提取目标图像的 SIFT 和 HOG 特征,接着将上述二维特征分别拓展为一维特征,最后将这两类特征衔接在一起形成一个新的特征来表征目标图像,从而实现 SIFT 和 HOG 特征在特征级的融合。

特征融合可以提高特征的表述能力,但其缺点也很明显:融合后的特征其维数明显高于融合前的特征维数。为了降低衔接在一起后的特征维数,本文又将 SIFT 和 HOG 融合后的特征做编码和池化操作以避免特征维数过高而引起的“维数诅咒”现象。编码的主要作用是用非线性映射将图像数据映射到另一个特征空间,以更好地表达原始图像的内容;而池化是一种特征加工过程,即对提取的特征进行二次加工降低特征维数,使提取的特征能够更好地用于后续的分类过程。

特征编码常用的方法有向量编码、Auto-encoders 压缩编码等。最近的研究^[10]表明可以用稀疏表示的方法实现特征的非线性映射。与传统方法相比,基于稀疏表示的编码方法能够提高特征表征能力。特征字典建立的流程如下:

假设字典为 $D=[D_1, D_2, \dots, D_K]$, 其中 D_i 是第 i 类的字典, $D_i=[d_1, d_2, \dots, d_L]$, 则有:

$$\min_{D_i, Z_i} \{ \|A_i - DZ_i\|_F^2 + \lambda \|Z_i\|_1 \}, \tag{1}$$

$$\text{s. t. } \|d_j\|_2 = 1, \forall j,$$

其中 A_i 是第 i 类训练样本, Z_i 是 A_i 在 D_i 上的表征矩阵。

文献^[10]中的方法也具有一定的局限性,比如在字典学习过程中并没有考虑词条与词条之间的内在联系,

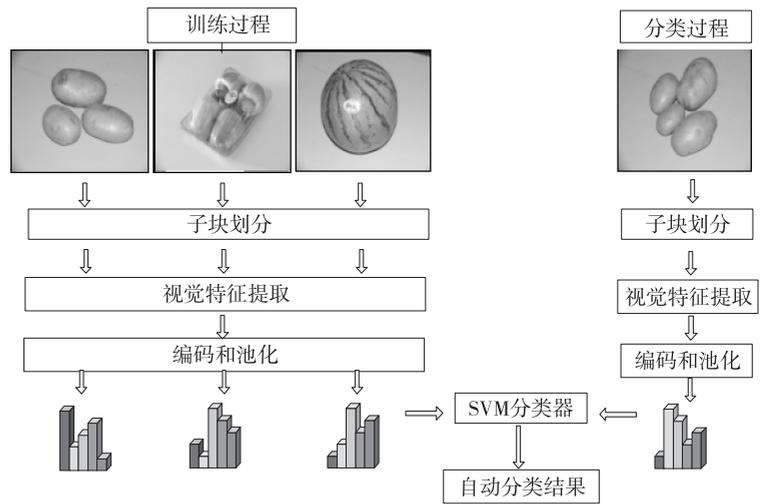


图 1 水果蔬菜自动分类方法流程图

Fig. 1 The flowchart of fruit and vegetable automatic classifications

这样的结果使得表征结果并不够精炼。为进一步提高非线性映射后特征的表征能力,本文采用文献[11]中的Fisher字典方法建立映射所需的字典,这样不仅可以用较少的原子来表示信号,还可以让这些重建后的特征具有较高的区分能力;同时还能获得对信号更为简洁的表示方式,从而更容易地获取信号中所蕴含的信息,具体方法如下:

$$\min_{(\mathbf{D}, \mathbf{X})} \left\{ \sum_{i=1}^K r(\mathbf{A}_i, \mathbf{D}, \mathbf{X}_i) + \lambda_1 \|\mathbf{X}\|_1 + \lambda_2 (\text{tr}(\mathbf{S}_W(\mathbf{X}) - \mathbf{S}_B(\mathbf{X})) + \eta \|\mathbf{X}\|_F^2) \right\}, \quad (2)$$

$$\text{s. t.} \quad \|d_n\|_2 = 1, \forall n,$$

其中 $\|\mathbf{X}\|_1$ 是稀疏度, $\text{tr}(\mathbf{S}_W(\mathbf{X}) - \mathbf{S}_B(\mathbf{X})) + \eta \|\mathbf{X}\|_F^2$ 是区分系数项, $r(\mathbf{A}_i, \mathbf{D}, \mathbf{X}_i)$ 表征的是数据的区分度,其定义为:

$$r(\mathbf{A}_i, \mathbf{D}, \mathbf{X}_i) = \|\mathbf{A}_i - \mathbf{D}\mathbf{X}_i\|_F^2 + \|\mathbf{A}_i - \mathbf{D}\mathbf{X}_i^j\|_F^2 + \sum_{j=1, j \neq i}^K \|\mathbf{D}_j \mathbf{X}_i^j\|_F^2. \quad (3)$$

通过固定一个变量求另外一个变量最优解的策略,循环迭代多次后就可以计算出公式(2)中的字典 \mathbf{D} 。

池化操作的结果是使得特征减少,同时池化还可以保持某种不变性(旋转、平移、伸缩等)。池化的数学定义就是找一个从 \mathbf{D}_i 映射到 \mathbf{P} 的函数 F , 即 $\mathbf{P} = F(\mathbf{D}_i)$ 。常用的有均值池化(Mean-pooling)、最大值池化(Max-pooling)和随机池化(Stochastic-pooling)3种,本文采用最大值池化,其定义为:

$$p_j = \max \{ |d_1|, |d_2|, \dots, |d_M| \}, \quad (4)$$

其中 p_j 是 \mathbf{P} 的第 j 个元素, M 是周围邻域元素的个数。

在分类过程中,用支持向量机(Support vector machine, SVM)^[12]分类器对输入图像做分类以识别水果蔬菜的种类。虽然分类器很多,但 SVM 分类器仍然具有分类精度较高、鲁棒性较好的特点,因此在分类阶段本文仍然采用 SVM 分类器。

2 实验

为公正客观地评价所提出方法的性能,本文在标准农产品超市数据库上^[1]将提出的方法与现有方法做了实验对比。

2.1 水果蔬菜数据库简介

该数据库包含了超市中常见的15类水果蔬菜产品,产品信息如下(括号中是对应英文名以及图片数量):亚加大马铃薯(Agata Potato, 201)、阿斯特里克斯马铃薯(Asterix Potato, 182)、腰果(Cashew, 210)、钻石桃子(Diamond Peach, 211)、富士苹果(Fuji Apple, 212)、澳洲青苹果(Granny-Smith Apple, 155)、哈密瓜(Honeydew Melon, 145)、猕猴桃(Kiwi, 171)、油桃(Nectarine, 247)、洋葱(Onion, 75)、橘子(Orange, 103)、李子(Plum, 264)、威廉斯梨(Williams Pear, 159)、大溪地柠檬(Taiti Lime, 106)以及西瓜(Watermelon, 192),总共2633张图片。

数据库中图片分辨率为 1024×768 像素,图片收集历时5个月,包含了不同形状和个数的图像,同类产品都是在不同时间收集所得。该数据库中农产品部分图例如图2所示。

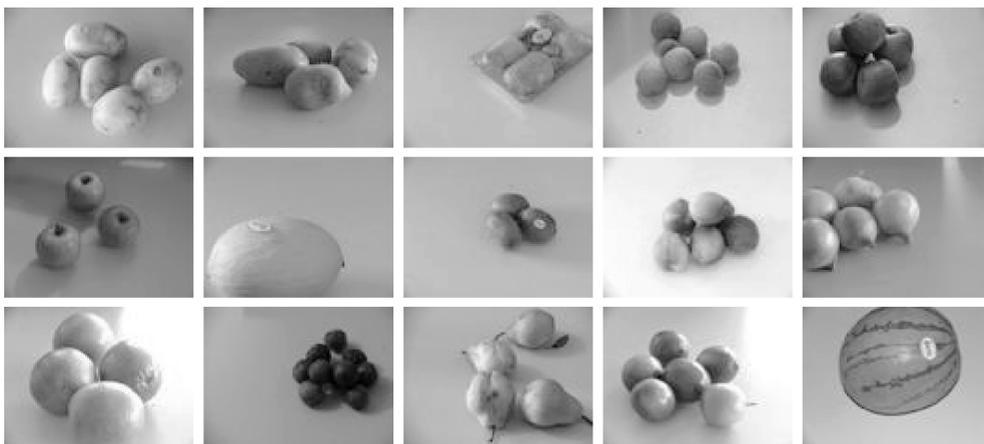


图2 数据库中水果蔬菜部分图片

Fig. 2 Samples of the fruit and vegetable database

2.2 参数设置

实验中为提高运算速度,将图像分辨率降低为原始图像的 0.1 倍,图像子块划分中滑动窗口大小为 16×16 像素,滑动步长为 6 个像素。SIFT 特征用文献[10]中的方法提取,池化方法中的参数也与文献[10]中的方法一致,HOG 特征用 Matlab 自带的函数计算而得。公式(2)中 $\lambda_1 = 0.005, \lambda_2 = 0.01$,字典更新次数为 10,其余参数与文献[11]完全相同。

2.3 实验结果比较

为了充分评价提出方法的识别效率,本文分别与文献[13]和文献[14]中的方法做了对比。文献[13]是继创建超市农产品数据库并提出相应自动识别方法^[1]之后,作者们对农产品自动识别提出的进一步解决方案。文献[14]对目标图像做小波变换后提取子带中的颜色、纹理等的统计特征来实现水果蔬菜的自动分类,其在农产品数据库^[1]的实验数据表明该方法具有较高的分类效率。

本文首先与文献[13]中方法做了对比,实验设置与文献[13]一致,即将每类农产品随机均匀划分为 5 份,其中 4 份用来训练,剩余 1 份作为测试数据,重复 5 次后求得平均识别率,实验对比结果如表 1 所示(文献[13]的实验结果来源于其文中的 Tab. 3)。

文献[13]提出的方法主要贡献在于将多种分类器的结果融合在一起实现自动分类。在该方法中尝试了多种解决方案,如 BOOST、SVM 等(各类方法详细介绍参见文献[13]),与文献[1]相比其识别率有了较大提高,如在 FSVM-PK-49 方法中,其识别率达到了 99.09%。

由于本文采用融合后的 SIFT 和 HOG 特征,借助于这两种方法的鲁棒性和有效性,再加之采用了编码和池化操作,所以与文献[13]中提出的方法相比,本文方法具有较高的识别率(99.44%)。出现这个结果的主要原因在于文献[13]方法只是尝试不同的分类

器并找出最佳的分类器组合方案,而本文方法不但利用相对具有较高分类性能的 SVM 分类器,更主要的一点在于提取出了鲁棒性较高的特征,并将这些特征融合在一起表征目标图像,因而具有较高的正确分类性能。

为进一步验证提出方法的有效性,本文又与文献[14]中的方法做了比较,实验设置与文献[14]中完全一致,即在每一类农产品中随机选择 50% 的图像用来训练,剩余 50% 的图像用作测试。随机测试共进行了 100 次,平均识别率如表 2 所示(文献[14]的实验结果来源于其文中的 Tab. 2)。

表 1 本文方法与文献[13]中方法识别率的比较

方法	识别率/%
MV-49	98.18
BOOST-49	97.65
BOOST-DEFAULT	95.82
BAGG-49	90.35
FSVM-PK-BEST	97.42
FSVM-PK-49	99.09
本文方法	99.44

表 2 本文方法与文献[14]中方法识别率的比较

Tab. 2 Comparisons between the proposed method and the method presented in [14]

测试集	颜色特征	纹理特征	颜色与纹理特征	本文方法
Agata Potato	56.43	74.25	95.04	94.76
Asterix Potato	52.74	65.93	90.10	95.96
Cashew	77.14	94.28	99.04	99.96
Diamond Peach	45.28	55.66	75.47	99.32
Fuji Apple	34.90	78.30	82.07	99.31
Granny-Smith Apple	30.76	89.74	96.15	95.87
Honeydew Melon	66.21	76.05	95.94	98.57
Kiwi	32.55	47.67	58.13	97.62
Nectarine	32.25	74.19	79.03	99.78
Onion	43.24	78.37	86.48	95.43
Orange	30.76	40.38	69.23	99.96
Plum	48.48	84.09	89.39	99.74

续表 2

测试集	颜色特征	纹理特征	颜色与纹理特征	本文方法
Spanish Pear	32.50	60.00	86.25	97.37
Taiti Lime	58.49	88.67	98.11	100
Watermelon	40.62	55.20	89.58	99.68
平均识别率	45.49	70.85	86.00	98.22

在文献[14]提出的方法中共采用了3种特征,即颜色特征、纹理特征以及颜色和纹理融合后的特征。从表中可以看出基于纹理特征的方法识别率高于基于颜色特征方法的识别率,而二者融合在一起的识别率又比单一特征的识别率高。

与文献[14]中的方法相比,本文提出的方法分类性能提升明显,主要原因是该农产品图像数据库采集时间比较长,采集条件比较复杂,使得同一类农产品的图片也有较大的颜色变化,因此颜色信息并不是特别稳定,而SIFT和HOG方法是在灰度图像上提取特征,因此提取出的特征不受颜色变化的影响,因而其识别率高于文献[14]中提出方法的识别率。

2.4 实验分析

本文提出的方法基于SIFT和HOG特征实现水果的自动分类,这两种方法具有较高的图像表征能力和鲁棒性,且这两类特征都是从灰度图像中提取而得,因此对颜色变化并不敏感;同时本文方法还采用了编码和池化实现了降低特征融合后的维数,从而提高特征表征能力的目的。综上,提出的方法具有较高的蔬菜水果自动分类性能,在标准农产品数据库^[1]上的实验数据都验证了提出方法的有效性。

3 结语

本文提出了一种超市中水果蔬菜的自动识别和分类的方法,在该方法中将SIFT和HOG特征融合在一起表征目标,然后通过编码和池化操作以提高特征的表征能力,最后用支持向量机实现自动分类。实验表明提出的方法具有较高的识别率,为实现超市中农产品的自动分类并进一步实现自动销售提供了技术和理论上的支持。

本文提出的方法也存在不足:在分类过程中仅考虑单一农产品的情形,即输入图片中只有某一类水果或蔬菜;然而在超市中农产品有时会混合在一起,在这种情形下提出的方法不能实现水果蔬菜的有效区分,因此提出水果蔬菜多类别自动分类的方法是未来研究工作的重点。

参考文献:

- [1] Rocha A, Hauagge D C, Wainer J, et al. Automatic fruit and vegetable classification from images[J]. Computers and Electronics in Agriculture, 2010, 70(1): 96-104.
- [2] 罗承成, 李书琴, 唐晶磊. 基于多示例学习的超市农产品图像识别[J]. 计算机应用, 2012, 32(6): 1560-1562.
Luo C C, Li S Q, Tang J J. Image recognition of agricultural products in supermarket based on multi-instance learning [J]. Journal of Computer Applications, 2012, 32(6): 1560-1562.
- [3] Biswas H, Hossain F. Automatic vegetable recognition system[J]. International Journal of Engineering Science Invention, 2013, 2(4): 37-41.
- [4] 陶华伟, 赵力, 奚吉, 等. 基于颜色及纹理特征的果蔬种类识别方法[J]. 农业工程学报, 2014, 30(16): 305-311.
Tao H W, Zhao L, Xi J, et al. Fruits and vegetables recognition based on color and texture features[J]. Transactions of the Chinese Society of Agricultural Engineering, 2014, 30(16): 305-311.
- [5] Dubey S R, Jalal S. Fruit and vegetable recognition by fusing colour and texture features of the image using machine learning[J]. International Journal of Applied Pattern Recognition, 2015, 2(2): 160-181.
- [6] Zhang Y D, Wang S H, Ji G L, et al. Fruit classification using computer vision and feed-forward neural network[J]. Journal of Food Engineering, 2014, 143(4): 167-177.
- [7] Gill J, Sandhu D P, Singh D P, et al. A review of automatic fruit classification using soft computing techniques[C]// International conference on computer, systems and electronics engineering. Johannesburg: ISAET Publications, 2014, 2: 99-105.
- [8] Lowe D G. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision,

- 2004, 60(2):91-110.
- [9] Dalal N, Triggs B. Histograms of oriented gradients for human detection[C]//International conference on computer vision and pattern recognition, San Diego: IEEE Publications, 2005, 1:886-893.
- [10] Yang J C, Yu K, Gong Y H, et al. Linear spatial pyramid matching using sparse coding for image classification [C]//International conference on computer vision and pattern recognition, Miami: IEEE Publications, 2009, 1: 1794-1801.
- [11] Yang M, Zhang L, Feng X C, et al. Sparse representation based fisher discrimination dictionary learning for image classification[J]. International Journal of Computer Vision, 2014, 109(3):209-232.
- [12] 王华秋, 王斌. 优化的邻近支持向量机在图像检索中的应用[J]. 重庆理工大学学报:自然科学版, 2014(9):66-71.
Wang H Q, Wang B. Application of optimized primal support vector machine in image retrieval[J]. Journal of Chongqing University of Technology: Natural Science, 2014(9):66-71.
- [13] Faria F A, Jefersson A, Santos D, et al. A framework for selection and fusion of pattern classifiers in multimedia recognition[J]. Pattern Recognition Letter, 2014, 39(1): 52-64.
- [14] Arivazhagan S, Newlin S R, Selva N S, et al. Fruit recognition using color and texture features[J]. Journal of Emerging Trends in Computing and Information Sciences, 2010, 1(2):90-94.

Fruit and Vegetable Automatic Classification Based on Appearance Feature

HE Jiangping, MA Yan, LI Qiang

(School of Information Engineering, Lanzhou University of Finance and Economics, Lanzhou 730020, China)

Abstract: This paper presents an automatic method of fruit and vegetable classification, which is based on appearance features to make contributions to automatic selling of these goods in supermarkets. First, the fruit and vegetable images are divided into overlapping regions, followed by the appearance features extraction, such as Scale-invariant feature transform and Histogram of Oriented Gradient. Second, the features are fused together for improving the representative ability. To reduce the feature dimension and enhance the discriminative power, the coding and pooling processes are performed on the fused features. Last, a Support Vector Machine classifier is designed based on the features and then the fruit and vegetable can be automatically classified and recognized. Compared to the state-of-the-art works, the presented method produces a higher recognition rate and provides a possible solution for the automatic selling of fruits and vegetables in supermarkets.

Key words: fruit and vegetable images; automatic classification; appearance features; coding; pooling

(责任编辑 游中胜)